

# DoF Analysis of the K-user MISO Broadcast Channel with Hybrid CSIT

Borzoo Rassouli\*, Chenxi Hao\* and Bruno Clerckx\*<sup>†</sup>

\*Communication and Signal Processing Group, Department of Electrical and Electronic Engineering  
Imperial College London, United Kingdom

<sup>†</sup>School of Electrical Engineering, Korea University, Korea  
Email: {b.rassouli12; chenxi.hao10; b.clerckx}@imperial.ac.uk

**Abstract**—We consider a  $K$ -user multiple-input single-output (MISO) broadcast channel (BC) where the channel state information (CSI) of user  $i$  ( $i = 1, 2, \dots, K$ ) may be either instantaneously perfect (P), delayed (D) or not known (N) at the transmitter with probabilities  $\lambda_P^i$ ,  $\lambda_D^i$  and  $\lambda_N^i$ , respectively. In this setting, according to the three possible CSIT for each user, knowledge of the joint CSIT of the  $K$  users could have at most  $3^K$  states. Although the results by Tandon et al. show that for the symmetric two user MISO BC (i.e.,  $\lambda_Q^i = \lambda_Q$ ,  $\forall i \in \{1, 2\}$ ,  $Q \in \{P, D, N\}$ ), the Degrees of Freedom (DoF) region depends only on the marginal probabilities, we show that this interesting result does not hold in general when  $K \geq 3$ . In other words, the DoF region is a function of all the joint probabilities. In this paper, given the marginal probabilities of CSIT, we derive an outer bound for the DoF region of the  $K$ -user MISO BC. Subsequently, we investigate the achievability of the outer bound in some scenarios. Finally, we show the dependence of the DoF region on the joint probabilities.<sup>1</sup>

## I. INTRODUCTION

In contrast to the point to point multiple-input multiple-output (MIMO) communication where the channel state information at the transmitter (CSIT) does not affect the multiplexing gain, in a multiple-input single-output (MISO) broadcast channel (BC), knowledge of CSIT is crucial for interference mitigation and beamforming purposes [1]. However, the assumption of perfect CSIT may not always be true in practice due to channel estimation error and feedback latency. Therefore, the idea of communication under some sort of imperfection in CSIT has gained more attention recently. The so called MAT algorithm was presented in [2] where it was shown that in terms of the degrees of freedom, even an outdated CSIT can result in significant performance improvement in comparison to the case with no CSIT. [3]–[6] investigate the time-correlated MISO BC where there is correlation between the feedback information and current channel state, while [7], [8] deal with the BC in a frequency-correlated setting. In [9] the synergistic benefits of alternating CSIT over fixed CSIT was presented in a two user MISO BC with two transmit antennas. In [10] and [11], the MISO BC with hybrid CSIT was considered, however our definition of hybrid CSIT is quite different with that of the aforementioned papers in the sense

that instead of having a fixed state (either P or D) for the CSIT of a particular user, the state is allowed to alternate among P, D and N as in [9].

Throughout the paper,  $f \sim o(\log P)$  is equivalent to  $\lim_{P \rightarrow \infty} \frac{f}{\log P} = 0$  and for a pair of integers  $m \leq q$ , the discrete interval is defined as  $[m : q] = \{m, m+1, \dots, q\}$ .  $Y_{[i:j]} = \{Y_i, Y_{i+1}, \dots, Y_j\}$ ,  $Y([i : j]) = \{Y(i), Y(i+1), \dots, Y(j)\}$  and  $Y^n = Y([1 : n])$ .

## II. SYSTEM MODEL

We consider a MISO BC, in which a base station with  $M$  antennas sends independent messages  $W_1, \dots, W_K$  to  $K$  single-antenna users ( $M \geq K$ ). In a flat fading scenario, the discrete-time baseband received signal of user  $k$  at channel use (henceforth, time instant)  $t$  can be written as

$$Y_k(t) = \mathbf{H}_k^H(t) \mathbf{X}(t) + W_k(t), \quad k \in [1 : K], \quad t \in [1 : n] \quad (1)$$

where  $\mathbf{X}(t) \in \mathbb{C}^M$  is the transmitted signal at time instant  $t$  satisfying the (per codeword) power constraint  $\sum_{t=1}^n \|\mathbf{x}(t)\|^2 \leq nP$ .  $W_k(t)$  and  $\mathbf{H}_k(t) \in \mathbb{C}^M$  are the additive noise and channel vector of user  $k$ , respectively, and are also assumed i.i.d. over the time instants and the users. We assume global perfect Channel State Information at Receivers (CSIR) and identity matrix for the covariance of the noise.

The rate tuple  $(R_1, R_2, \dots, R_K)$ , in which  $R_i = \frac{\log(|W_i|)}{n}$ , is achievable if there exists a coding scheme such that the probability of error in decoding  $W_i$  at user  $i$  ( $i \in [1 : K]$ ) can be made arbitrarily small with sufficiently large coding block length. The DoF region is defined as  $\{(d_1, \dots, d_K) | \exists (R_1, R_2, \dots, R_K) \in C(P) \text{ such that } d_i = \lim_{P \rightarrow \infty} \frac{R_i}{\log P}, \forall i\}$  where  $C(P)$  is the capacity region (i.e., the closure of the set of achievable rate tuples).

The hybrid CSIT model means that at some time instants the transmitter has a Perfect (P) instantaneous knowledge of the CSI of a particular user, whereas at some time instants it receives the CSI with Delay (D) and finally, at some time instants the CSI of the user is Not known (N) at the transmitter. When there is delayed CSIT, we assume that the feedback delay is larger than the coherence time of the channel making the feedback information completely independent of the current channel state. In this configuration, the joint CSIT of all the  $K$  users has at most  $3^K$  states. For example, in a 3 user

<sup>1</sup>This work was partially supported by the Seventh Framework Programme for Research of the European Commission under grant number HARP-318489.

	Time Slots		
Users	D	N	P
	P	D	N
	P	P	P

Fig. 1: A CSIT pattern with  $\lambda_{DDP} = \lambda_{NDP} = \lambda_{PNP} = \frac{1}{3}$

MISO BC, they will be  $PPP, PDN, \dots$  with corresponding probabilities  $\lambda_{PPP}, \lambda_{PDN}, \dots$  and the marginal probability of perfect CSIT for user 1 is  $\lambda_P^1 = \sum_{Q, Q' \in \{P, D, N\}} \lambda_{PQQ'}$ .

By CSIT pattern we refer to the knowledge of CSIT represented in a space-time matrix where the rows and columns represent users and time slots, respectively. Figure 1 shows an example of a CSIT pattern, in which the transmitter knows the channels of users 2 and 3 perfectly at time slot 1 and has no information about the channel of user 1. The CSI of user 1 will be known in the next time instants (i.e., time slot 2) due to feedback delay and is completely independent of the channel in time slot 2.

The main result of this paper is that given the marginal probabilities of CSIT, an outer bound for the DoF region is provided regardless of the CSIT pattern. Further, through a simple example, we show that the DoF region for  $K \geq 3$  must be a function of the CSIT pattern rather than only of the marginal probabilities in contrast to the results of [9] for the 2-user case. This dependency is equivalent to having the optimal DoF region as a function of  $\lambda_{DDP}, \lambda_{PDN}, \dots$  in such a way that they do not add up to produce only the marginal probabilities.

### III. MAIN RESULTS

**Theorem.** Let  $\pi^j(\cdot)$  be an arbitrary permutation of size  $j$  over the indices  $(1, 2, \dots, K)$ , and  $\alpha_{\pi^j}(\cdot)$  be a permutation of  $\pi^j$  satisfying

$$(\lambda_P^{\alpha_{\pi^j}(i)} + \lambda_D^{\alpha_{\pi^j}(i)}) \leq (\lambda_P^{\alpha_{\pi^j}(i+1)} + \lambda_D^{\alpha_{\pi^j}(i+1)}), \quad i \in [1 : j-1]. \quad (2)$$

Given the marginal probabilities of CSIT for user  $i$  (which can be any two of  $\lambda_P^i, \lambda_D^i$  and  $\lambda_N^i$ , since  $\lambda_P^i + \lambda_D^i + \lambda_N^i = 1$ ), an outer bound for the DoF region of the  $K$ -user MISO BC with  $M$  transmit antennas at the transmitter ( $M \geq K$ ) is defined by the following sets of inequalities

$$\sum_{i=1}^j \frac{d_{\pi^j(i)}}{i} \leq 1 + \sum_{i=2}^j \frac{\sum_{r=1}^{i-1} \lambda_P^{\pi^j(r)}}{i(i-1)} \quad (3)$$

$$\sum_{i=1}^j d_{\pi^j(i)} \leq 1 + \sum_{i=1}^{j-1} (\lambda_P^{\alpha_{\pi^j}(i)} + \lambda_D^{\alpha_{\pi^j}(i)}) \quad (4)$$

for all  $\pi^j$  and  $j \in [1 : K]$ . For the symmetric scenario, the sets of inequalities are simplified as

$$\sum_{i=1}^j \frac{d_{\pi^j(i)}}{i} \leq 1 + \lambda_P \sum_{i=2}^j \frac{1}{i} \quad (5)$$

$$\sum_{i=1}^j d_{\pi^j(i)} \leq 1 + (j-1)(\lambda_P + \lambda_D). \quad (6)$$

For  $K = 2$ , the outer bound boils down to the optimal DoF region in [9].

### IV. PROOF OF THEOREM

For simplicity, we assume  $j = K$ , since it is obvious that each subset of users with cardinality  $j$  ( $j < K$ ) can be regarded as a  $j$ -user BC. Also, we assume the identity permutation (i.e.,  $\pi^K(i) = i$ ) while the results could be easily applied to any other arbitrary permutation.

A. *Proof of  $\sum_{i=1}^K \frac{d_i}{i} \leq 1 + \sum_{i=2}^K \frac{\sum_{r=1}^{i-1} \lambda_P^r}{i(i-1)}$*

First, we improve the channel by giving the message and observation of user  $i$  to users  $[i+1 : K]$  ( $i \in [1 : K-1]$ ). Hence, from Fano's inequality,

$$nR_i \leq I(W_i; Y_{[1:i]}^n | W_{[1:i-1]}^n, \Omega^n) + n\epsilon_n \quad (7)$$

where  $\Omega^n$  denotes the global CSIR up to time instant  $n$ ,  $W_0 = \emptyset$  and  $\epsilon_n$  goes to zero as  $n$  goes to infinity. This improvement results in a degraded broadcast channel [12]. Therefore, according to [13], since feedback does not increase the capacity of degraded broadcast channels, we can ignore the delayed CSIT (D) and replace them with No CSIT (N). Therefore, it is equivalent to having the channel of user  $i$  perfectly known with probability  $\lambda_P^i$  and not known otherwise. From now on, we ignore the term  $n\epsilon_n$  for simplicity (since later it will be divided by  $n$  and  $n \rightarrow \infty$ ) and write

$$\sum_{i=1}^K \frac{nR_i}{i} \leq \sum_{i=1}^K \frac{I(W_i; Y_{[1:i]}^n | W_{[1:i-1]}^n, \Omega^n)}{i} \quad (8)$$

$$\leq h(Y_1^n | \Omega^n) + \sum_{i=2}^K \left[ \frac{h(Y_{[1:i]}^n | W_{[1:i-1]}^n, \Omega^n)}{i} - \frac{h(Y_{[1:i-1]}^n | W_{[1:i-1]}^n, \Omega^n)}{i-1} \right] + no(\log P) \quad (9)$$

where  $Y_0 = \emptyset$  and we have used the fact that

$$\frac{h(Y_{[1:K]}^n | W_{[1:K]}^n, \Omega^n)}{nK} \sim o(\log P).$$

since with the knowledge of  $W_{[1:K]}$  and  $\Omega^n$ , the observations  $Y_{[1:K]}^n$  can be reconstructed within the noise distortion. From the chain rule of entropies, each of the terms in the summation in (9) can be written as

$$\sum_{t=1}^n \left[ \frac{h(Y_{[1:i]}(t) | W_{[1:i-1]}^t, Y_{[1:i]}^{t-1}, \Omega^t)}{i} - \frac{h(Y_{[1:i-1]}(t) | W_{[1:i-1]}^t, Y_{[1:i-1]}^{t-1}, \Omega^t)}{i-1} \right]. \quad (10)$$

By adding  $Y_i^{t-1}$  in the conditioning of the second entropy, (10) will be increased. Therefore,

$$\sum_{i=1}^K \frac{nR_i}{i} \leq \underbrace{h(Y_1^n | \Omega^n)}_{\leq n \log P} + \sum_{i=2}^K \sum_{t=1}^n \left[ \frac{h(Y_{[1:i]}(t) | U_{i,t}, \Omega(t))}{i} - \frac{h(Y_{[1:i-1]}(t) | U_{i,t}, \Omega(t))}{i-1} \right] + no(\log P) \quad (11)$$

where  $U_{i,t} = (W_{[1:i-1]}, Y_{[1:i]}^{t-1}, \Omega^{t-1})$  and  $\Omega(t)$  is the global CSIR at time instant  $t$ . Before going further, the following lemma is needed.

**Lemma 1.** Let  $\Gamma_N = \{Y_1, Y_2, \dots, Y_N\}$  be a set of  $N (\geq 2)$  arbitrary random variables and  $\Psi_i^j(\Gamma_N)$  be a sliding window of size  $j$  over  $\Gamma_N$  ( $1 \leq i, j \leq N$ ) starting from  $Y_i$  i.e.,

$$\Psi_i^j(\Gamma_N) = Y_{(i-1)N+1}, Y_{iN+1}, \dots, Y_{(i+j-2)N+1}$$

where  $(\cdot)_N$  defines the modulo  $N$  operation. Then,

$$(N-1)h(Y_{[1:N]}|A) \leq \sum_{i=1}^N h(\Psi_i^{N-1}(\Gamma_N)|A) \quad (12)$$

where  $A$  is an arbitrary condition.

*Proof:* We prove the lemma by induction. It is obvious that (12) holds for  $N = 2$ . In other words,  $h(Y_{[1:2]}|A) \leq \sum_{i=1}^2 h(Y_i|A)$ . Now, considering that (12) is valid for  $N (\geq 2)$ , we show that it also holds for  $N+1$ . Replacing  $N$  with  $N+1$ , we have

$$\begin{aligned} Nh(Y_{[1:N+1]}|A) &= h(Y_{[1:N+1]}|A) + (N-1)h(Y_{[1:N-1]}, \overbrace{Y_N, Y_{N+1}}^Z | A) \\ &\leq h(Y_{[1:N+1]}|A) + \sum_{i=1}^N h(\Psi_i^{N-1}(\Phi_N)|A) \\ &= h(Y_{[1:N+1]}|A) + h(\Psi_1^{N-1}(\Phi_N)|A) \\ &\quad + \sum_{i=2}^N h(\Psi_i^N(\Gamma_{N+1})|A) \\ &= h(Y_N | Y_{N+1}, Y_{[1:N-1]}, A) \\ &\quad + h(\Psi_1^{N-1}(\Phi_N)|A) + h(Y_{N+1}, Y_{[1:N-1]}|A) \\ &\quad + \sum_{i=2}^N h(\Psi_i^N(\Gamma_{N+1})|A) \\ &= h(Y_N | Y_{N+1}, Y_{[1:N-1]}, A) \\ &\quad + h(\Psi_1^{N-1}(\Phi_N)|A) + \sum_{i=2}^{N+1} h(\Psi_i^N(\Gamma_{N+1})|A) \\ &= h(Y_N | Y_{N+1}, Y_{[1:N-1]}, A) \end{aligned} \quad (13)$$

(14)

(15)

$$+ h(Y_{[1:N-1]}|A) + \sum_{i=2}^{N+1} h(\Psi_i^N(\Gamma_{N+1})|A) \quad (16)$$

$$\leq h(Y_N | Y_{[1:N-1]}, A) + h(Y_{[1:N-1]}|A) + \sum_{i=2}^{N+1} h(\Psi_i^N(\Gamma_{N+1})|A) \quad (17)$$

$$= h(\Psi_1^N(\Gamma_{N+1})|A) + \sum_{i=2}^{N+1} h(\Psi_i^N(\Gamma_{N+1})|A) = \sum_{i=1}^{N+1} h(\Psi_i^N(\Gamma_{N+1})|A) \quad (18)$$

where in (13),  $\Phi_N = \{Y_{[1:N-1]}, Z\}$  and we have used the validity of (12) for  $N$ . In (14), we have used the fact that  $\Psi_i^N(\Gamma_{N+1}) = \Psi_i^{N-1}(\Phi_N)$  for  $i \in [2 : N]$ . In (15), the chain rule of entropies is used and in (16), the sliding window is written in terms of its elements. Finally, in (17), the fact that conditioning does not increase the differential entropy is used. Therefore, since (12) is valid for  $N = 2$  and from its validity for  $N (\geq 2)$  we could show it also holds for  $N+1$ , the proof is complete. ■

Each term in the summation of (11) can be rewritten as

$$\frac{(i-1)h(Y_{[1:i]}(t) | U_{i,t}, \Omega(t)) - ih(Y_{[1:i-1]}(t) | U_{i,t}, \Omega(t))}{i(i-1)} \quad (19)$$

$$\leq \frac{\sum_{r=1}^i [h(\Psi_r^{i-1}(\Gamma_i) | U_{i,t}, \Omega(t)) - h(Y_{[1:i-1]}(t) | U_{i,t}, \Omega(t))]}{i(i-1)} \quad (20)$$

$$= \frac{\sum_{r=1}^{i-1} [h(Y_i(t) | E_{r,i}, U_{i,t}, \Omega(t)) - h(Y_r(t) | E_{r,i}, U_{i,t}, \Omega(t))]}{i(i-1)} \quad (21)$$

where  $\Gamma_i = \{Y_{[1:i]}(t)\}$ ,  $E_{r,i} = \{Y_{[1:i-1]}(t)\} - \{Y_r(t)\}$ , (20) is from the application of lemma 1 and (21) is from the chain rule of entropies. Before going further, the following lemma is needed.

**Lemma 2.** In the  $K$ -user MISO BC defined in (1), for the users  $m, q \in [1 : K]$  ( $m \neq q$ ), we have

$$\lim_{P \rightarrow \infty} \frac{h(Y_m(t)|A) - h(Y_q(t)|A)}{\log P} \leq \begin{cases} 1 & \text{CSIT of } \mathbf{H}_q(t) \text{ is } P \\ 0 & \text{CSIT of } \mathbf{H}_q(t) \text{ is } N \end{cases} \quad (22)$$

where  $A$  is a condition such as the condition of entropies in (21) or later in (32). Interestingly, (22) is only a function of the CSIT of the second user.

*Proof:* Based on the four possible states for the joint CSIT of  $\mathbf{H}_m(t)$  and  $\mathbf{H}_q(t)$ , we have

1) CSIT of  $\mathbf{H}_m(t)$  is  $N$  or  $P$  and CSIT of  $\mathbf{H}_q(t)$  is  $P$ :

$$h(Y_m(t)|A) - h(Y_q(t)|A) \leq \underbrace{h(Y_m(t)|A)}_{\leq \log P} - \underbrace{h(Y_q(t)|A, W_{[1:K]})}_{o(\log P)} \quad (23)$$

A Gaussian input with the conditional covariance matrix of  $\Sigma_{X|A} = P\mathbf{u}_q^\perp \mathbf{u}_q^{\perp H}$  achieves the upper bound, where  $\mathbf{u}_q^\perp$  is a unit vector in the direction orthogonal to  $\mathbf{H}_q(t)$  (since  $\mathbf{H}_q(t)$  is known).

2) CSIT of  $\mathbf{H}_m(t)$  is  $N$  and CSIT of  $\mathbf{H}_q(t)$  is  $N$ : In this case both  $Y_m(t)$  and  $Y_q(t)$  are statistically equivalent (i.e., having the same probability density functions, and subsequently, the same entropies.) Therefore,

$$h(Y_m(t)|A) - h(Y_q(t)|A) = 0 \quad (24)$$

3) CSIT of  $\mathbf{H}_m(t)$  is  $P$  and CSIT of  $\mathbf{H}_q(t)$  is  $N$ : This is shown in [14].

From (11) and (21), we have

$$\sum_{i=1}^K \frac{nR_i}{i} \leq \sum_{i=2}^K \sum_{t=1}^n \sum_{r=1}^{i-1} \frac{h(Y_i(t)|A_{r,i,t}) - h(Y_r(t)|A_{r,i,t})}{i(i-1)} \quad (25)$$

$$+ n \log P + no(\log P) \quad (26)$$

$$\leq n \log P + \sum_{i=2}^K \sum_{r=1}^{i-1} \frac{n\lambda_P^r}{i(i-1)} \log P + no(\log P) \quad (27)$$

where  $A_{r,i,t}$  in the conditioning of the entropies in (21) and (27) is from the application of lemma 2 and the fact that  $n$  is sufficiently large. Therefore,

$$\sum_{i=1}^K \frac{d_i}{i} \leq 1 + \sum_{i=2}^K \frac{\sum_{r=1}^{i-1} \lambda_P^r}{i(i-1)}. \quad (28)$$

It is obvious that the same approach can be applied to any other permutation on  $(1, 2, \dots, K)$  which results in (3).

**B. Proof of  $\sum_{i=1}^K d_i \leq 1 + \sum_{i=1}^{K-1} (\lambda_P^{\alpha_{\pi^K}(i)} + \lambda_D^{\alpha_{\pi^K}(i)})$**

We enhance the channel in two ways:

- 1) Like the approach in [9], whenever there is delayed CSIT ( $D$ ), we assume that it is perfect instantaneous CSIT ( $P$ ), but we keep the probability of delayed CSIT. In other words, the CSIT of user  $i$  is perfect with probability  $\lambda_P^i + \lambda_D^i$  and unknown otherwise.
- 2) We give the message of user  $i$  to users  $[i+1 : K]$ .

Therefore,

$$nR_i \leq I(W_i; Y_i^n | W_{[1:i-1]}, \Omega^n) + n\epsilon_n, \quad \forall i \in [1 : K]. \quad (29)$$

By summing (29) over users and writing the mutual information in terms of differential entropies,

$$\sum_{i=1}^K nR_i \leq \overbrace{h(Y_1^n | \Omega^n)}^{\leq n \log P} + \sum_{i=2}^K [h(Y_i^n | W_{[1:i-1]}, \Omega^n) - h(Y_{i-1}^n | W_{[1:i-1]}, \Omega^n)] + no(\log P). \quad (30)$$

By the application of *Csiszár sum identity* [15], the term in the summation could be written as

$$\sum_{t=1}^n [h(Y_i(t) | F_{i,t}, \Omega(t)) - h(Y_{i-1}(t) | F_{i,t}, \Omega(t))] \quad (31)$$

where

$$F_{i,t} = (W_{[1:i-1]}, \Omega^{t-1}, Y_{i-1}^{t-1}, Y_i([t+1 : n])).$$

Therefore,

$$\sum_{i=1}^K nR_i \leq n \log P + \sum_{i=2}^K \sum_{t=1}^n [h(Y_i(t) | F_{i,t}, \Omega(t)) - h(Y_{i-1}(t) | F_{i,t}, \Omega(t))] \quad (32)$$

■ and finally, by applying the results of lemma 2 to (32), we have

$$\sum_{i=1}^K d_i \leq 1 + \sum_{i=2}^K (\lambda_P^{i-1} + \lambda_D^{i-1}) = 1 + \sum_{i=1}^{K-1} (\lambda_P^i + \lambda_D^i). \quad (33)$$

Since the same approach holds for any arbitrary permutation of size  $K$  on  $(1, \dots, K)$ , we have

$$\sum_{i=1}^K d_i \leq 1 + \min_{\pi^K(\cdot)} \sum_{i=1}^{K-1} (\lambda_P^{\pi^K(i)} + \lambda_D^{\pi^K(i)}) \quad (34)$$

and it is obvious that  $\alpha_{\pi^K}(\cdot)$  will minimize (34) if it satisfies (2) (for  $j = K$ ).

## V. ACHIEVABILITY

In this section, we consider the achievability of the symmetric case. For  $K \geq 3$ , we show that given the marginal probabilities of CSIT, there exists at least one CSIT pattern that achieves the outer bound in some scenarios. We investigate the following two scenarios:

**A.  $\lambda_D = 0$**

In this case,  $2^K - 1$  inequalities are active and the remaining inequalities become inactive and the region is defined by  $2^K - 1$  hyperplanes in  $R_+^K$  that has the following  $K$  corner points:

$$(1, \lambda_P, \dots, \lambda_P), (\lambda_P, 1, \lambda_P, \dots, \lambda_P), \dots, (\lambda_P, \dots, \lambda_P, 1)$$

The corner points are simply achieved by a scheme that has  $N$  time slots and consists of two parts: in the first  $\lambda_P N$  time slots, zero forcing beamforming (ZFBF) is carried out where each user receives one interference-free symbol. In the remaining  $\lambda_N N$  time slots, only one particular user (depending on the corner point of interest) is scheduled.

**B.  $\lambda_N \leq \frac{\lambda_P}{\sum_{j=2}^K \frac{1}{j}}$**

Before going further, we need the following simple lemma.

**Lemma 3.** The minimum probability of delayed CSIT for sending order- $j$  symbols in the  $K$ -user MAT is

$$\lambda_D^{\min}(K, j) = 1 - \frac{K - j + 1}{K \sum_{i=j}^K \frac{1}{i}}. \quad (35)$$

*Proof:* From [2], the MAT algorithm is based on a concatenation of  $K$  phases. Phase  $j$  takes  $(K-j+1)\binom{K}{j}$  order- $j$  messages as its input, takes  $\binom{K}{j}$  time slots and produces  $j\binom{K}{j+1}$  order- $j+1$  messages as its output. In each time slot of phase  $j$ , the transmitter sends a random linear combination of the  $(K-j+1)$  symbols to a subset  $S$  of receivers,  $|S| = j$ . Sending the overheard interferences from the

remaining  $(K - j)$  receivers to receivers in subset  $S$  enables them to successfully decode their  $(K - j + 1)$  symbols by constructing a set of  $(K - j + 1)$  linearly independent equations. Therefore, the transmitter needs to know the channel of only  $(K - j)$  receivers. In other words, at each time slot of phase  $j$ , the feedback of  $(K - j)$  CSI is enough. In the MAT algorithm the number of output symbols that phase  $j$  produces should match the number of input symbols of phase  $j + 1$ . The ratio between the input of phase  $j + 1$  and output of phase  $j$  is:

$$\frac{(K - j) \binom{K}{j+1}}{j \binom{K}{j}} = \frac{(K - j)}{j}.$$

This means that  $(K - j)$  repetition of phase  $j$  will produce the inputs needed by  $j$  repetition of phase  $j + 1$ . In general, in order to have an integer number for repetitions, we multiply phase 1 by  $K!$  (i.e., repeat it  $K!$  times), phase 2 by  $\frac{K!}{(K-1)}$ , and so on. Therefore, phase  $j$  will be repeated  $((j-1)!(K-j)!)K \binom{K}{j}$  times which takes  $((j-1)!(K-j)!)K \binom{K}{j}$  time slots. Since  $(K-j)$  feedbacks from each time slot is sufficient, the number of feedbacks will be  $((j-1)!(K-j)!)K \binom{K}{j} (K-j)$ . For a successive decoding or order- $j$  symbols, all the higher order symbols must be decoded successfully. Therefore, instead of having delayed CSIT at all time instants from all users, the minimum probability of delayed CSIT is the number of feedbacks from phase  $j$  to  $K$  divided by the whole number of time slots multiplied by the number of users,

$$\begin{aligned} \lambda_D^{min}(K, j) &= \frac{\sum_{i=j}^K (i-1)!(K-i)!K \binom{K}{i} (K-i)}{\sum_{i=j}^K (i-1)!(K-i)!K \binom{K}{i} K} \\ &= 1 - \frac{K-j+1}{K \sum_{i=j}^K \frac{1}{i}}. \end{aligned}$$

In this case (i.e.,  $\lambda_N \leq \frac{\lambda_D}{\sum_{j=2}^K \frac{1}{j}}$ ), The region has  $2^K - 1$  corner points. In other words, if the coordinates of a point are shown as  $(p_1, p_2, \dots, p_K)$ , there are  $\binom{K}{j}$  ( $j \in [1 : K]$ ) points where  $j$  of its  $K$  coordinates are  $\frac{1+\lambda_P \sum_{i=2}^j \frac{1}{i}}{\sum_{i=1}^j \frac{1}{i}}$  and the remaining  $K-j$  coordinates are  $\lambda_P$ . The achievable scheme is based on a concatenation of ZFBF and MAT as follows. For the  $\binom{K}{j}$  corner points, we write

$$\lambda_P = \frac{M_1}{N_1}, \lambda_D = \frac{M_2}{N_2}, \lambda_D^{min}(j, 1) = \frac{m}{n} \quad (36)$$

where  $m, n, M_i$  and  $N_i$  ( $i = 1, 2$ ) are integers. Making a common denominator between  $\lambda_P$  and  $\lambda_D$  we have

$$\lambda_P = \frac{nM_1N_2}{nN_1N_2}, \lambda_D = \frac{nN_1M_2}{nN_1N_2}. \quad (37)$$

We construct  $nN_1N_2$  time slots where the CSIT of each user can be Perfect (P) or Delayed (D) in  $nM_1N_2$  or  $nN_1M_2$  time slots, respectively. In the first  $nM_1N_2$  time slots, ZFBF is carried out. In the remaining  $n(N_1N_2 - M_1N_2)$  time slots,  $j$ -user MAT algorithm is done. At each time slot of the ZFBF part, 1 interference-free symbol is received by each user and in the MAT part,  $\frac{n(N_1N_2 - M_1N_2)}{1 + \frac{1}{2} + \dots + \frac{1}{j}}$  symbols are sent to each of

P	N	N
P	N	N
P	N	N

(a)

P	N	N
N	P	N
N	N	P

(b)

Fig. 2: Two symmetric CSIT patterns having the same marginal probabilities (i.e.,  $\lambda_P = 1 - \lambda_N = \frac{1}{3}$ .)

the users in subset  $S$  (with  $|S| = j$ ) where  $S$  depends on the corner point of interest. In order to do the MAT algorithm in the second part, the minimum probability of delayed CSIT should be met

$$nN_1M_2 \geq \lambda_D^{min}(j)n(N_1N_2 - M_1N_2) \quad (38)$$

Dividing both sides by  $nN_1N_2$ ,

$$\lambda_D \geq \lambda_D^{min}(j, 1)(1 - \lambda_P) = \lambda_D^{min}(j, 1)(\lambda_D + \lambda_N) \quad (39)$$

which results in

$$\lambda_N \leq \frac{\lambda_D}{\sum_{i=2}^j \frac{1}{i}}. \quad (40)$$

Since it should be valid for all  $j$ , we have

$$\lambda_N \leq \frac{\lambda_D}{\sum_{i=2}^K \frac{1}{i}}. \quad (41)$$

## VI. DEPENDENCY OF THE DOF REGION ON THE CSIT PATTERN

Here, we show that two different CSIT patterns, though having the same marginal probabilities, do not necessarily have the same DoF region. Consider the two simple symmetric CSIT patterns shown in figure 2. According to the theorem, the DoF region of both has an outer bound with the corner points  $(1, \frac{1}{3}, \frac{1}{3})$ ,  $(\frac{1}{3}, 1, \frac{1}{3})$  and  $(\frac{1}{3}, \frac{1}{3}, 1)$ . It is obvious that the corner points are achievable for pattern (a), and in what follows we show that they are not achievable for pattern (b). We write,

$$nR_1 \leq I(W_1; Y_1^n | \Omega^n) \quad (42)$$

$$nR_1 \leq I(W_1; Y_1^n | \Omega^n, W_2) \quad (43)$$

Adding (42) and (43) results in

$$2nR_1 \leq I(W_1; Y_1^n | \Omega^n) + I(W_1; Y_1^n | \Omega^n, W_2). \quad (44)$$

By doing the same for  $R_2$ , we have

$$2nR_2 \leq I(W_2; Y_2^n | \Omega^n) + I(W_2; Y_2^n | \Omega^n, W_1). \quad (45)$$

Finally, the rate of user 3 is written as

$$nR_3 \leq I(W_3; Y_3^n | \Omega^n, W_1, W_2). \quad (46)$$

Therefore,

$$\begin{aligned}
& 2nR_1 + 2nR_2 + nR_3 \\
& \leq \underbrace{+h(Y_2^n|\Omega^n, W_1) - h(Y_1^n|\Omega^n, W_1)}_{\leq \frac{n}{3} \log P} + \underbrace{h(Y_3^n|\Omega^n, W_1, W_2)}_{\leq \frac{n}{3} \log P} \\
& \quad + \underbrace{+h(Y_1^n|\Omega^n, W_2) - h(Y_2^n|\Omega^n, W_2)}_{\leq \frac{n}{3} \log P} + \underbrace{h(Y_1^n|\Omega^n)}_{\leq n \log P} + \underbrace{h(Y_2^n|\Omega^n)}_{\leq n \log P} \\
& \quad - \underbrace{h(Y_1^n|\Omega^n, W_1, W_2) - h(Y_2^n|\Omega^n, W_1, W_2)}_{\leq -h(Y_1^n, Y_2^n|\Omega^n, W_1, W_2)} \\
& \quad - h(Y_3^n|\Omega^n, W_1, W_2, W_3) \quad (47) \\
& \leq \frac{8n}{3} \log P + h(Y_3^n|\Omega^n, W_1, W_2) - h(Y_1^n, Y_2^n|\Omega^n, W_1, W_2) \\
& = \frac{8n}{3} \log P + \underbrace{h(Y_3^n|T_n) - h(Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n|T_n)}_{o(\log P)} \\
& \quad - \underbrace{h(Y_{1,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n|T_n, Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n)}_{\sim o(\log P)} \\
& \leq -h(Y_{1,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n|T_n, Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n, W_3) \sim o(\log P) \quad (48) \\
& \leq \frac{8n}{3} \log P \quad (49)
\end{aligned}$$

where  $T_n = \{\Omega^n, W_1, W_2\}$  and in (47), the difference terms are first written as a time summation of instantaneous differences, as in (31). Then, lemma 2 is applied to the differences resulting in the values written under the braces. We have split the observation of users 1 and 2 in terms of the joint CSIT, i.e.,  $Y_1^n = (Y_{1,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n)$  and  $Y_2^n = (Y_{2,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n)$ . The first  $o(\log P)$  in (48) is due to the fact that there is at least one unknown CSIT (N) in the joint states of user 1 and user 2 (i.e., PN, NP and NN. see rows 1 and 2 of the CSIT pattern shown in figure 2 (b)) and the fact that (22) is upper bounded by zero when the CSIT of the second term is N. Therefore, for pattern (b), the following inequalities hold which make its DoF region inside that of pattern (a):

$$\begin{aligned}
2d_1 + 2d_2 + d_3 & \leq \frac{8}{3} \\
2d_1 + d_2 + 2d_3 & \leq \frac{8}{3} \\
d_1 + 2d_2 + 2d_3 & \leq \frac{8}{3}. \quad (50)
\end{aligned}$$

Motivated by this simple example, we can have the following set of inequalities for the 3-user MISO BC

$$\begin{aligned}
2d_1 + 2d_2 + d_3 & \leq 2 + (\lambda_P^1 + \lambda_D^1) + (\lambda_P^2 + \lambda_D^2) \\
& \quad + (\lambda_{PP-} + \lambda_{PD-} + \lambda_{DP-} + \lambda_{DD-}) \\
2d_1 + d_2 + 2d_3 & \leq 2 + (\lambda_P^1 + \lambda_D^1) + (\lambda_P^3 + \lambda_D^3) \\
& \quad + (\lambda_{P-P} + \lambda_{P-D} + \lambda_{D-P} + \lambda_{D-D}) \\
d_1 + 2d_2 + 2d_3 & \leq 2 + (\lambda_P^2 + \lambda_D^2) + (\lambda_P^3 + \lambda_D^3) \\
& \quad + (\lambda_{-PP} + \lambda_{-PD} + \lambda_{-DP} + \lambda_{-DD}) \quad (51)
\end{aligned}$$

where a dash in the above means that the CSIT of the corresponding user is not important (for example,  $\lambda_{PD-} =$

$\lambda_{PDP} + \lambda_{PDD} + \lambda_{PDN}$  which is a summation over all the possible values for the CSIT of user 3). The same approach could be easily extended to the  $K$ -user MISO BC which is omitted for brevity. It is obvious that none of the above inequalities can have its right-hand side written in terms of only marginal probabilities. Therefore, in contrast to the two user scenario, marginal probabilities of CSIT are not sufficient for defining the DoF region of the general  $K$ -user MISO BC, and having the same marginal probabilities does not guarantee the same DoF region.

## VII. CONCLUSION

Given the marginal probabilities of CSIT, an outer bound was derived for the DoF region of the  $K$ -user MISO BC with hybrid CSIT alternating. This outer bound was shown to be achievable by specific CSIT patterns in certain regions. Through an example, we showed that in general, the DoF region of the  $K$ -user MISO BC (when  $K \geq 3$ ) is a function of CSIT patterns or equivalently the  $3^K$  state probabilities rather than the sole marginal probabilities.

## REFERENCES

- [1] B. Clerckx and C. Oestges, *MIMO Wireless Networks, 2nd Edition*. Academic Press, 2013.
- [2] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, 2012.
- [3] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated miso broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, 2013.
- [4] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Comm. Letters*, vol. 16, no. 7, pp. 1084–1087, July 2012.
- [5] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," vol. arxiv/1205.3474, May, 2012.
- [6] P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the K-user time correlated broadcast channel with delayed CSIT," in *ISIT 2013, IEEE International Symposium on Information Theory, July 7-12, 2013, Istanbul, Turkey*, Istanbul, TURKEY, July 2013. [Online]. Available: <http://www.eurecom.fr/publication/4000>
- [7] C. Hao and B. Clerckx, "Imperfect and unmatched CSIT is still useful for the frequency correlated MISO broadcast channel," in *IEEE ICC, Budapest, Hungary, Jun. 2013*, available on arXiv:1302.6521.
- [8] —, "MISO broadcast channel with imperfect and (un)matched CSIT in the frequency domain: DoF region and transmission strategies," in *IEEE PIMRC*, Sept. 2013.
- [9] R. Tandon, S. Jafar, S. Shamaï Shitz, and H. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, 2013.
- [10] K. Mohanty and M. Varanasi, "On the DoF region of the K-user MISO broadcast channel with hybrid CSIT," vol. available on arXiv:1312.1309, Dec. 2013.
- [11] S. Amuru, R. Tandon, and S. Shamaï, "On the degrees-of-freedom of the 3-user MISO broadcast channel with hybrid CSIT," in *IEEE ISIT*, 2014, pp. 2137–2141.
- [12] T. M. Cover and J. A. Thomas, *Elements of Information Theory, second edition*. New York: Wiley-Interscience, 2006.
- [13] A. Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379 – 381, may 1978.
- [14] A. Davoodi and S. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamaï and Wigger on the collapse of degrees of freedom under finite precision CSIT," *Arxiv* <http://arxiv.org/abs/1403.1541>.
- [15] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.